United States Patent [19]

Davis

[11] **Patent Number: 4,961,152**

[45] **Date of Patent: Oct. 2, 1990**

[54] **ADAPTIVE COMPUTING SYSTEM**

[75] Inventor: **Lawrence Davis**, Cambridge, Mass.

[73] Assignee: **Bolt Beranek and Newman Inc.**, Cambridge, Mass.

[21] Appl. No.: **205,154**

[22] Filed: **Jun. 10, 1988**

[51] **Int. Cl.**[5] **G06F 15/18**; G05B 13/00

[52] **U.S. Cl.** **364/513**; 364/274; 364/916; 364/972; 364/221.1; 364/200; 364/900

[58] **Field of Search** ..... 364/513, 200 MS File, 364/274, 221.1, 900 MS File, 916, 972; 382/14, 15; 365/49

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,697,242	9/1987	Holland et al.	364/513
4,821,333	4/1989	Gillies	382/41

OTHER PUBLICATIONS

"Classifier System Learning of a Boolean Function", Stewart W. Wilson, The Rowland Institute for Science, Research Memo RIS No. 27R, 2/9/86.

"Knowledge Growth in an Artificial Animal", Stewart W. Wilson, Rowland Institute for Science.

"Classifier Systems and the Animal Problem", Stewart W. Wilson, The Rowland Institute for Science.

"A critical Review of Genetic Algorithms", J. David Schaffer and John J. Grefenstette.

"Improving the Performance of Genetic Algorithms in Classifier Systems", L. B. Booker, Proc. Intern. Conf. on Gen. Alg. and their Appl. Pittsburg, PA, Jul. 1985, pp. 80–92.

"Classifier Systems with Hamming Weights", L. Davis and D. K. Young, 3–31–88, pp. 1–12.

Primary Examiner—Jerry Smith

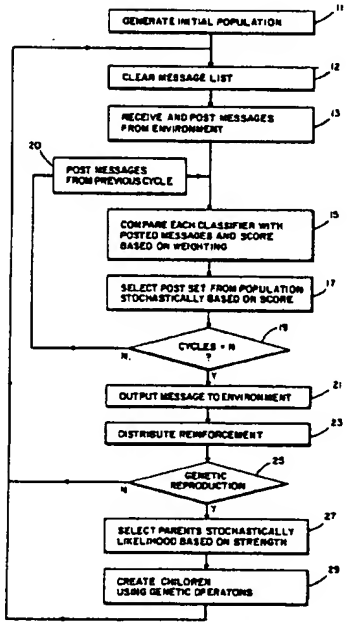
Assistant Examiner—Paul Gordon

Attorney, Agent, or Firm—Henry D. Pahl, Jr.

[57] **ABSTRACT**

The adaptive computing system disclosed herein employs a data structure involving a multiplicity of classifiers, each of which includes a match string of characters which the system attempts to match up with messages generated either by the environment or by other classifiers. Associated with each match string is a response message definer or action part which defines the response to be given when a match is obtained. To facilitate matching in a noisy or changing environment, there is associated with each character in the match string a weighting value and the degree of match is judged by means of a score value which is a function of the weighting values of the characters which match between the match string and the message.

10 Claims, 2 Drawing Sheets



US-PAT-NO: 4961152

DOCUMENT-IDENTIFIER: US 4961152 A
See image for Certificate of Correction

TITLE: Adaptive computing system

DATE-ISSUED: October 2, 1990

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Davis; Lawrence	Cambridge	MA	N/A	N/A

US-CL-CURRENT: 706/13

ABSTRACT:

The adaptive computing system disclosed herein employs a data structure involving a multiplicity of classifiers, each of which includes a match string of characters which the system attempts to match up with messages generated either by the environment or by other classifiers. Associated with each match string is a response message definer or action part which defines the response to be given when a match is obtained. To facilitate matching in a noisy or changing environment, there is associated with each character in the match string a weighting value and the degree of match is judged by means of a score value which is a function of the weighting values of the characters which match between the match string and the message.

10 Claims, 3 Drawing figures

Exemplary Claim Number: 6

Number of Drawing Sheets: 2

---------- KWIC ---------

Brief Summary Text - BSTX (13):

The match string for each classifier is compared to each input signal message and a corresponding score value is generated which is a function of the weighting values of the matching bits. The score values generated by the different classifiers are compared and respective response messages are output for classifiers having the better scores. The strength value for each classifier whose output message was appropriate is modified as a function of the respective reinforcement signal. Periodically, genetic operations are applied to generate new classifiers based upon the modification of selected ones of the match strings and selected sets of weighting values, the selection process including a randomizing component and a component based on the respective strength values.

Detailed Description Text - DETX (9):

In accordance with the present invention, genetic operations are performed on the weight values associated with each of the bits in the match string as well as upon the match string itself. The process is represented in FIG. 1 by the block designated by reference character 29. The genetic operations performed on the weights may be performed in conjunction with those on the match string or may be performed independently thereof. FIG. 3 represents the application of a powerful crossover genetic operator applied in coordinated fashion to both the match string and the associated weight values. In this example, parent 1 comprises the match string 010101 and parent 2 comprises a match string which is all zeroes. A crossover point is determined at random. In the example illustrated in FIG. 3, the crossover point is assumed to be between the second and third bit positions. It should be understood that multiple crossover points, also chosen at random, may be usefully implemented particularly with longer match strings. With reference to FIG. 2, the combining with crossover yields two children of the parents, one with the match string 010000 and the other with the match string 000101. In each case in the illustrated example, the weight value associated with each bit of the child string is taken from the same parent which produced the match bit. An alternative, however, would be to average the weights from the two parents for each bit position. As is understood by those skilled in the art, this sort of crossover genetic operation provides a chance for the generation of new classifiers which combine useful aspects of two different parent classifiers.

As indicated previously, the selection of parents is performed as a stochastic process with the likelihood of selection being based on the relative strength of the various classifiers. On the other hand, since the stochastic process includes a random component, a healthy diversity in the overall population is maintained.

Claims Text - CLTX (9):
comparing the score values generated by the different classifiers;

Claims Text - CLTX (24):
comparing the score values generated by the different classifiers;

United States Patent [19]

Cooper et al.

[11] **Patent Number: 5,054,093**

[45] **Date of Patent: Oct. 1, 1991**

[54] **PARALLEL, MULTI-UNIT, ADAPTIVE, NONLINEAR PATTERN CLASS SEPARATOR AND IDENTIFIER**

[76] Inventors: **Leon N. Cooper**, 49 Intervale Rd.; **Charles Elbaum**, 85 Lorraine Ave.; **Douglas L. Reilly**, 69 E. Manning St.; **Christopher L. Scofield**, 478 Morris Ave., all of Providence, R.I. 02906

[21] Appl. No.: **179,427**

[22] Filed: **Apr. 8, 1988**

Related U.S. Application Data

[60] Division of Ser. No. 775,144, Sep. 12, 1985, Pat. No. 4,760,604, which is a continuation-in-part of Ser. No. 702,188, Feb. 15, 1985, abandoned.

[51] **Int. Cl.**[5] **G06K 9/00**

[52] **U.S. Cl.** **382/14**; 382/36

[58] **Field of Search** 382/14, 15, 16; 364/513

[56] **References Cited**

U.S. PATENT DOCUMENTS

3,333,248	7/1967	Greenberg et al.	382/15
3,601,802	8/1971	Nakagome et al.	382/37
4,030,068	6/1977	Banz	382/38
4,451,929	5/1984	Yoshida	382/15
4,479,241	10/1984	Buckley	382/15
4,503,557	3/1985	Maeda	382/34
4,566,123	1/1986	Yoshida	382/15

OTHER PUBLICATIONS

B. G. Batchelor: "Practical Approach to Pattern Classification", Plenum Press, London and New York; (1974).

B. G. Batchelor: "Pattern Recognition, Ideas in Practice", Plenum Press, London and New York; (1978).

K. Udagawa et al: "A Parallel Two-Stage Decision Method for Statistical Character Recognition . . . ", Electronics and Communications in Japan (1965).

J. Schurmann, "Zur Zeichen und Worterkennung beim Automatischen Anschriftenlesen", Wissenschaftlichl. Berichte, vol. 52, No. 1/2 (1979).

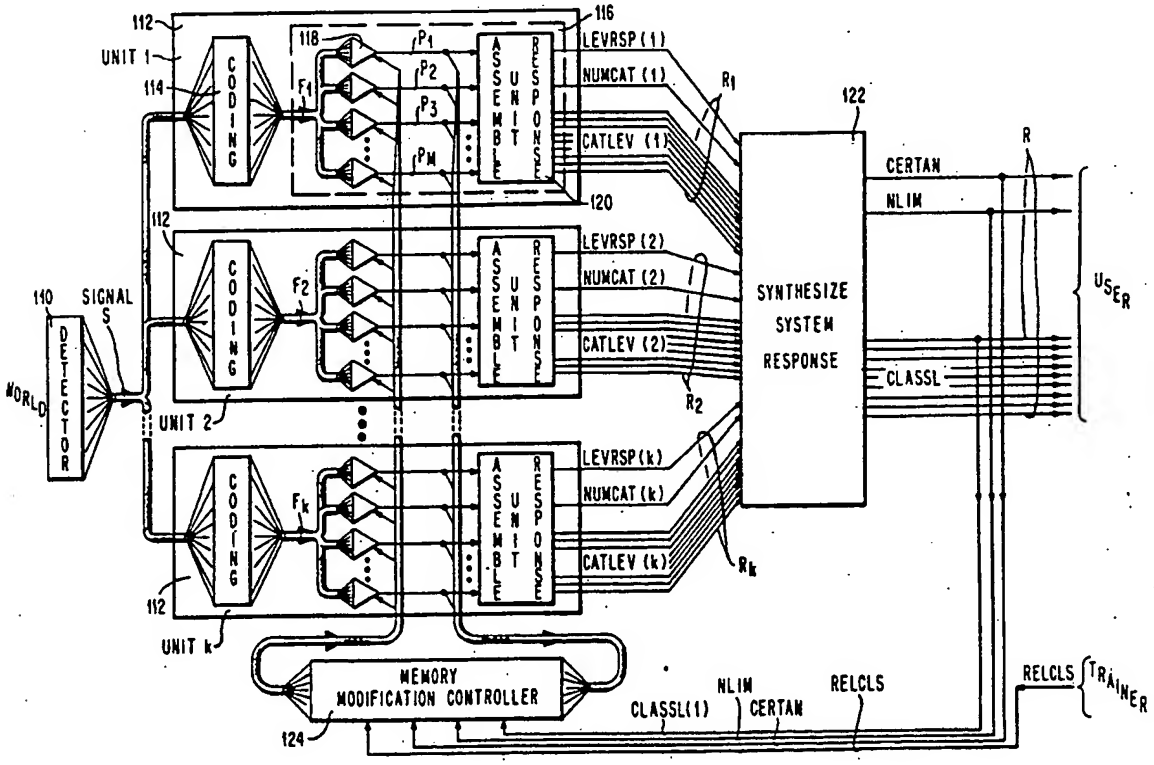
Primary Examiner—Leo H. Boudreau

Attorney, Agent, or Firm—Sprung Horn Kramer & Woods

[57] **ABSTRACT**

A system for separating and identifying classes of patterns or events which are not necessarily linearly separable. During a training process of the system, new prototypes are created when prior stored prototypes fail to properly classify an input pattern, and previously stored prototypes are modified when an input pattern falls within the sphere of influence of a prototype associated with a different class than the input pattern.

70 Claims, 8 Drawing Sheets



US-PAT-NO: 5054093

DOCUMENT-IDENTIFIER: US 5054093 A

TITLE: Parallel, multi-unit, adaptive, nonlinear pattern class separator and identifier

DATE-ISSUED: October 1, 1991

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Cooper; Leon N.	Providence	RI	02906	N/A
Elbaum; Charles	Providence	RI	02906	N/A
Reilly; Douglas L.	Providence	RI	02906	N/A
Scofield; Christopher L.	Providence	RI	02906	N/A

US-CL-CURRENT: 382/159, 382/224

ABSTRACT:

A system for separating and identifying classes of patterns or events which are not necessarily linearly separable. During a training process of the system, new prototypes are created when prior stored prototypes fail to properly classify an input pattern, and previously stored prototypes are modified when an input pattern falls within the sphere of influence of a prototype associated with a different class than the input pattern.

70 Claims, 9 Drawing figures

Exemplary Claim Number: 1

Number of Drawing Sheets: 8

---------- KWIC ---------

Brief Summary Text - BSTX (25):

Similarly, although only a single pattern feature encoding device may be provided, the provision of a multiplicity of pattern classifiers, looking at a different aspect of the encoding device output and each trained in a slightly different way, makes it possible to separate the pattern into one of a large number of possible classes. It will be appreciated that when a very large number of possible classes are involved, the pattern classifier is able to distinguish between subtle differences in the patterns which it "sees".

Claims Text - CLTX (68):

52. The pattern classification method defined in claim 51, wherein said label of said given prototype is converted from a first phase type to a second phase type if said scalar distance thereof is reduced below a prescribed minimum value.

US005438629A

United States Patent [19]

Moed et al.

[11] Patent Number: 5,438,629

[45] Date of Patent: Aug. 1, 1995

[54] METHOD AND APPARATUS FOR INPUT CLASSIFICATION USING NON-SPHERICAL NEURONS

[75] Inventors: Michael C. Moed, Norwalk; Chih-Ping Lee, Danbury, both of Conn.

[73] Assignee: United Parcel Service of America, Inc., Atlanta, Ga.

[21] Appl. No.: 901,429

[22] Filed: Jun. 19, 1992

[51] Int. Cl.[6] G06K 9/00

[52] U.S. Cl. 382/156; 382/159; 395/21

[58] Field of Search 382/14, 15, 36, 49; 395/21, 22, 23; G06K 9/62

[56] References Cited

U.S. PATENT DOCUMENTS

3,701,095	10/1972	Yamaguchi	382/49
3,950,733	4/1976	Cooper et al.	395/24
4,044,243	8/1977	Cooper et al.	395/20
4,326,259	4/1982	Cooper et al.	364/715.11
4,599,693	7/1986	Denenberg	395/77
4,954,963	9/1990	Penz et al.	364/513
5,010,512	4/1991	Hartstein et al.	364/807
5,033,006	7/1991	Ishizuka et al.	364/513
5,054,093	10/1991	Cooper et al.	382/14
5,058,180	10/1991	Kahn	382/14
5,060,278	10/1991	Fukumizu	382/14
5,067,164	11/1991	Denker et al.	382/15
5,086,479	2/1992	Takenaga et al.	382/14
5,119,438	6/1992	Ueda et al.	382/14
5,181,171	1/1993	McCormack et al.	364/421
5,214,715	5/1993	Carpenter et al.	382/14
5,214,744	5/1993	Schweizer et al.	395/21
5,218,646	6/1993	Sirat et al.	382/14
5,239,594	8/1993	Yoda	382/15
5,245,697	9/1993	Suzuoka	395/22
5,247,584	9/1993	Krogmann	382/14
5,260,871	11/1993	Goldberg	364/413.02
5,265,192	11/1993	McCormack	382/22

FOREIGN PATENT DOCUMENTS

0133589	6/1987	Japan .
0033687	2/1990	Japan .
0238588	9/1990	Japan .
0242392	9/1990	Japan .
0300876	12/1990	Japan .

OTHER PUBLICATIONS

Authors: Sheldon Gruber and Leda Villalobos, title "Neural Network Based Inspection of Machined Surfaces Using Laser Scattering", SPIE vol. 1265 Industrial Inspection II (1990); pp. 85–93.

1990 IEEE Transactions on Systems, Man, and Cybernetics, vol. 20, No. 4, Jul./Aug. 1990, pp. 816–825, authors Yuzo Hirai and Yasuyuki Tsukui, Title "Position Independent Pattern Matching by Neural Network".

(List continued on next page.)

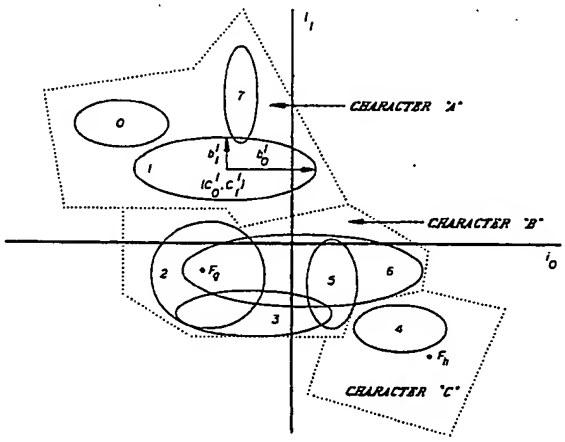
Primary Examiner—Yon J. Couso

Attorney, Agent, or Firm—William H. Murray; Steve Mendelsohn

[57] ABSTRACT

A classification method and apparatus for classifying an input into one of a plurality of possible outputs. Information representative of the input is compared to a neuron, where the neuron comprises a boundary defined by two or more neuron axes of different length. One of the possible outputs is then selected as corresponding to the input in accordance with that comparison. The invention is also a training method and apparatus for creating a new neuron or adjusting an existing neuron. A feature vector representative of a training input is generated, where the training input corresponds to one of a plurality of possible outputs. If no existing neuron corresponding to the training input encompasses the feature vector, then a new neuron is created, where the new neuron comprises a boundary defined by two or more neuron axes of different length. If the neuron encompasses the feature vector and if the neuron does not correspond to the training input, then the neuron is adjusted spatially, where the adjusted neuron comprises a boundary defined by two or more adjusted neuron axes of different length.

76 Claims, 10 Drawing Sheets



US-PAT-NO: 5438629

DOCUMENT-IDENTIFIER: US 5438629 A
See image for Certificate of Correction

TITLE: Method and apparatus for input classification using non-spherical neurons

DATE-ISSUED: August 1, 1995

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Moed; Michael C.	Norwalk	CT	N/A	N/A
Lee; Chih-Ping	Danbury	CT	N/A	N/A

US-CL-CURRENT: 382/156, 382/159 , 706/20 , 706/25

ABSTRACT:

A classification method and apparatus for classifying an input into one of a plurality of possible outputs. Information representative of the input is compared to a neuron, where the neuron comprises a boundary defined by two or more neuron axes of different length. One of the possible outputs is then selected as corresponding to the input in accordance with that comparison. The invention is also a training method and apparatus for creating a new neuron or adjusting an existing neuron. A feature vector representative of a training input is generated, where the training input corresponds to one of a plurality of possible outputs. If no existing neuron corresponding to the training input encompasses the feature vector, then a new neuron is created, where the new neuron comprises a boundary defined by two or more neuron axes of different length. If the neuron encompasses the feature vector and if the neuron does not correspond to the training input, then the neuron is adjusted spatially, where the adjusted neuron comprises a boundary defined by two or more adjusted neuron axes of different length.

76 Claims, 13 Drawing figures

Exemplary Claim Number: 1

Number of Drawing Sheets: 10

---------- KWIC ---------

Detailed Description Text - DETX (87):
In one preferred embodiment, each cluster classifier may implement a classification algorithm unique to that character cluster, or shared by only a subset of the total number of character clusters. Each cluster classifier may therefore employ neurons that exist in a feature space unique to that character cluster. For example, training for the "P", "R", "B" cluster may employ a particular set of Grid features, while training for the "O", "C", "D", "U", "Q" cluster may employ a different set of Hadamard features. In that case, different training procedures are performed for each different cluster classifier, where only inputs corresponding to those characters of the associated cluster are used for each different training procedure.

Detailed Description Text - DETX (93):
In an alternative embodiment, two or more different classifiers may be arranged in parallel. In such case, a voting scheme may be employed to select the appropriate output by comparing the outputs of each different classifier.

Claims Text - CLTX (21):
10. The classification method of claim 1, wherein said first number is equal to the number of said selected neurons that are associated with said first possible output, and said second number that is equal to the number of said selected neurons that are associated with said second possible output.

US005761383A

United States Patent [19]

Engel et al.

[11] Patent Number: 5,761,383

[45] Date of Patent: Jun. 2, 1998

[54] **ADAPTIVE FILTERING NEURAL NETWORK CLASSIFIER**

[75] Inventors: **Stephen J. Engel**, E. Northport; **Dennis Buckland**, Bethpage, both of N.Y.

[73] Assignee: **Northrop Grumman Corporation**, Los Angeles, Calif.

[21] Appl. No.: **867,719**

[22] Filed: **Jun. 3, 1997**

Related U.S. Application Data

[63] Continuation of Ser. No. 430,529, Apr. 27, 1995, abandoned.

[51] **Int. Cl.**^{6} **G06F 15/18**

[52] **U.S. Cl.** **395/21**; 395/22; 395/23

[58] **Field of Search** 395/21, 22, 23; 382/190

[56] **References Cited**

U.S. PATENT DOCUMENTS

5,003,490	3/1991	Castelaz et al.	395/513
5,092,343	3/1992	Spitzer et al.	128/733
5,133,021	7/1992	Carpenter et al.	382/15
5,165,069	11/1992	Vitt et al.	358/335
5,263,097	11/1993	Katz et al.	382/190
5,386,689	2/1995	Bozich et al.	60/39.33
5,402,520	3/1995	Schmitta	395/22
5,467,428	11/1995	Ulug	395/23
5,479,572	12/1995	Marcantonio	395/22
5,533,383	7/1996	Greene et al.	73/40.5 A
5,579,232	11/1996	Tong et al.	364/474.17
5,586,220	12/1996	Spiker	395/22
5,649,065	7/1997	Lo et al.	395/23
5,680,627	10/1997	Anglea et al.	395/751

OTHER PUBLICATIONS

Widrow et al., Neural Nets for adaptive filtering and adaptive pattern recognition, Computer, IEEE 1988, pp. 25–39, Mar. 1988.

Broomhead et al., A parallel architecture for nonlinear adaptive filtering and pattern recognition, IEEE conf on artificial Neural network, 1989, pp. 265–269.

Hill et al., Antenna beamforming for EW using adaptive layered networks, IEE Colloquium on 'Signal processing in Electronic Warfare' (Digest #1994/025), pp. 2/1–5, Jan. 31, 1994.

Dony et al., Optimally integrated adaptive learning, 1993 ICASSP, pp. 609–12, Apr. 30, 1993.

Blacknell et al., Optimum classification of non-Gaussian processes using neural networks, IEEE proceedings, pp. 55–66, Feb. 1994.

Ko et al., Neural network based novelty filtering for signal detection enhancement, 1992 IEEE Midwest Symposium, pp. 252–255.

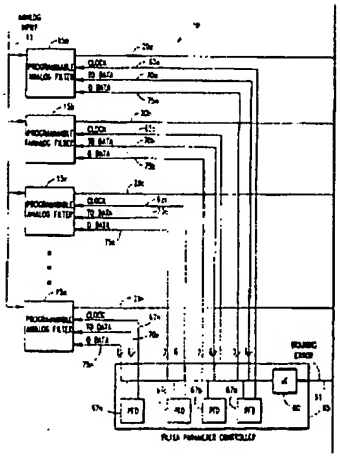
Lippman (1987) "An Introduction to Computing with Nueral Nets".

Primary Examiner—Robert W. Downs
Assistant Examiner—Sanjiv Shah
Attorney, Agent, or Firm—Terry J. Anderson; Karl J. Hoch, Jr.

[57] **ABSTRACT**

An adaptive filtering neural network classifier for classifying input signals, includes a neural network and one or more adaptive filters for receiving input analog signals to be classified and generates inputs for the classifier. Each adaptive filter is characterized as having a predetermined number of operating parameters. An analog to digital converter converts each input signal into a digital signal before input to the neural network. The neural network processes each digital signal to generate therefrom a plurality of weighted output signals in accordance with the type of network implemented. One of the weighted output signals represents a class for the input signal, and an error signal representing a difference between the weighted output signals and a predetermined desired output is also generated by the network. A control device responsive to the error signal generates a further set of operating filter parameters for input to each of the adaptive filters to change the operating response thereof and minimize the error signal.

14 Claims, 5 Drawing Sheets



1st stage computes
orderly scores & outputs
These scores is parallel to
the next stage

US006507843B1

(12) **United States Patent**
Dong et al.

(10) **Patent No.: US 6,507,843 B1**
(45) **Date of Patent: Jan. 14, 2003**

(54) **METHOD AND APPARATUS FOR CLASSIFICATION OF DATA BY AGGREGATING EMERGING PATTERNS**

(75) Inventors: **Guozhu Dong**, Beavercreek, OH (US); **Jinyan Li**, Victoria (AU); **Limsoon Wong**, Kuala Lumpur (MY); **Xiuzhen Zhang**, Victoria (AU)

(73) Assignee: **Kent Ridge Digital Labs**, Singapore (SG)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **09/492,753**

(22) Filed: **Jan. 27, 2000**

(30) **Foreign Application Priority Data**

Aug. 14, 1999 (SG) 9903916
Aug. 16, 1999 (SG) 9903958

(51) **Int. Cl.**[7] **G06F 17/30**
(52) **U.S. Cl.** **707/6**; 707/102; 707/5; 706/50
(58) **Field of Search** 707/6, 5, 3, 2, 707/104.1, 102, 1, 10, 101, 4, 100, 50; 706/50, 12; 705/14, 1, 26; 709/224; 435/6; 379/88.22; 600/300

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,832,182 A * 11/1998 Zhang et al. 706/50

OTHER PUBLICATIONS

Guozhu Dong, Department of CSE, Wright State Univ. and Jinyan Li, Department of CSSE, The University of Melbourne, "Efficient Mining of Emerging Patterns Discovering Trends and Differences". Aug. 1, 1999.

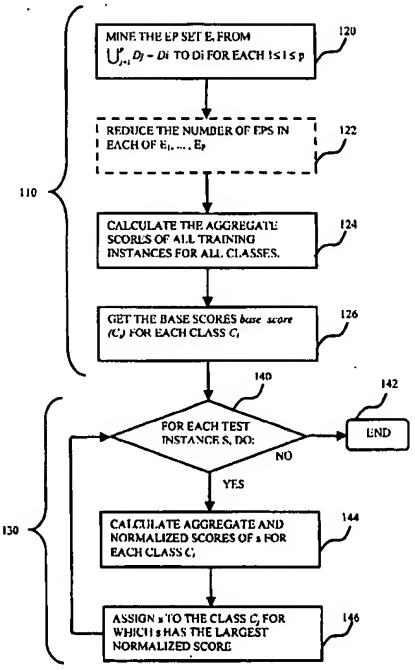
* cited by examiner

Primary Examiner—Sanjiv Shah
Assistant Examiner—Gwen Liang
(74) *Attorney, Agent, or Firm*—Nath & Associates PLLC; Harold L. Novick; Marvin C. Berkowitz

(57) **ABSTRACT**

Emerging patterns (EPs) are itemsets having supports that change significantly from one dataset to another. A classifier, CAEP, is disclosed using the following main ideas based on EPs: (i) Each EP can sharply differentiate the class membership of a (possibly small) fraction of instances containing the EP, due to the big difference between the EP's supports in the opposing classes; the differentiating power of the EP is defined in terms of the EP's supports and ratio, on instances containing the EP. (ii) For each instance t, by aggregating (**124**) the differentiating power of a fixed, automatically selected set of EPs, a score is obtained for each class (**126**). The scores for all classes are normalized (**144**) and the largest score determines t's class (**146**). CAEP is suitable for many applications, even those with large volumes of high dimensional data. CAEP does not depend on dimension reduction on data and is usually equally accurate on all classes even if their populations are unbalanced.

32 Claims, 3 Drawing Sheets



aggregate scores of the instances are calculated for all of the classes

US006549646B1

(12) **United States Patent**
Yeh et al.

(10) **Patent No.: US 6,549,646 B1**
(45) **Date of Patent: Apr. 15, 2003**

not for input text

(54) **DIVIDE-AND-CONQUER METHOD AND SYSTEM FOR THE DETECTION OF LUNG NODULE IN RADIOLOGICAL IMAGES**

(75) Inventors: **Hwa-Young Michael Yeh**, Potomac, MD (US); **Jyh-Shyan Lin**, Potomac, MD (US); **Yuan-Ming Fleming Lure**, Potomac, MD (US); **Xin-Wei Xu**, Gaithersburg, MD (US); **Ruiping Li**, Rockville, MD (US); **Rong Feng Zhuang**, Hyattsville, MD (US)

(73) Assignee: **Deus Technologies, LLC**, Rockville, MD (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **09/503,840**

(22) Filed: **Feb. 15, 2000**

(51) **Int. Cl.**[7] **G06K 9/00**

(52) **U.S. Cl.** **382/132**; 382/157; 382/199; 382/203; 382/205; 382/257; 382/274

(58) **Field of Search** 382/130, 132, 382/156, 172, 203, 224, 157, 199, 257, 274, 204, 205; 128/922, 925

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,907,156	A	3/1990	Doi et al.	
5,463,548	A	10/1995	Asada et al.	
5,638,458	A *	6/1997	Giger et al.	382/132
6,335,980	B1 *	1/2002	Armato et al.	382/132

OTHER PUBLICATIONS

Y.S.P. Chiou, Y.M.F. Lure, Hybrid Lung Nodule Detection (HLLND) system, *Cancer Letter 77*(1994), pp. 119–126.

McNitt–Gray et al., Feature Selection in the Pattern Classification Problem of Digital Chest Radiograph Segmentation, *Transaction on Medical Imaging*, vol. 14, No. 3 (9/95), pp. 537–547.

Duryea et al., A fully automated algorithm for the segmentation of lung fields on digital chest radiographic images, *Med. Phys. vol. 22 No. 2* (2/95) pp 183–191.

Armato, III et al., Computerized detection of abnormal asymmetry in digital chest radiographs[a], *Med. Phys. 21 vol. 21, No. 11*, Nov. 1994, pp. 1761–1768.

Maria J. Carreira et al., Automatic Segmentation of Lung Fields on Chest Radiographic Images, *Computers and Biomedical Research 32*, (1999) pp. 283–303.

Neal F. Vittitoe et al. Identification of lung regions in chest radiographs using Markov random field modeling., *Med. Phys., vol. 25, No. 6* (6/98), pp. 976–985.

Akira Hasegawa et al., A Shift–Invariant Neural Network for the Lung Field Segmentation in Chest Radiography, *Journal of VLSI Signal Processing 18* (1998) pp. 241–250.

Osamu Tsujii et al., Automated segmentation of anatomic regions in chest radiographs using an adaptive–sized hybrid neural network, *Med. Phys., vol. 25, No. 6* (Jun. 1998), pp. 998–1007.

(List continued on next page.)

Primary Examiner—Jon Chang
Assistant Examiner—Charles Kim
(74) *Attorney, Agent, or Firm*—Venable, LLP; Jeffrey W. Gluck; Robert Kinberg

(57) **ABSTRACT**

A divide-and-conquer (DAC) method and system improve the detection of abnormalities, like lung nodules, in radiological images via the use of zone-based digital image processing and artificial neural networks. The DAC method and system divide the lung zone into different zones in order to enhance the efficiency in detection. Different image enhancement techniques are used for each different zone to enhance nodule images, as are different zone-specific techniques for selecting suspected abnormalities, extracting image features corresponding to selected abnormalities, and classifying the abnormalities as either true or false abnormalities.

102 Claims, 11 Drawing Sheets

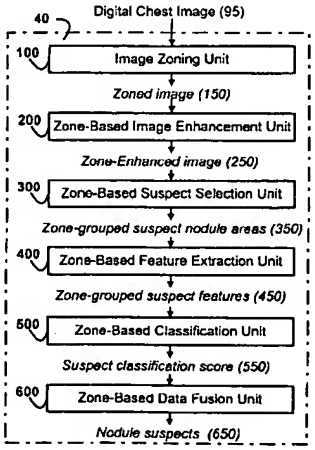


Fig 9
classification score generated by different classifiers of different zones

US 20020099730A1

(19) United States

(12) Patent Application Publication (10) Pub. No.: US 2002/0099730 A1

Brown et al. (43) Pub. Date: Jul. 25, 2002

(54) AUTOMATIC TEXT CLASSIFICATION SYSTEM

(75) Inventors: Daniel Brown, London (GB); Benjamin Anthony Janes, Surrey (GB); Murray Steele, London (GB); Richard James Cooper, London (GB)

Correspondence Address:
FRISHAUF, HOLTZ, GOODMAN & LANGER & CHICK, PC
767 THIRD AVENUE
25TH FLOOR
NEW YORK, NY 10017-2023 (US)

(73) Assignee: Applied Psychology Research Limited, London (GB)

(21) Appl. No.: 09/854,838

(22) Filed: May 14, 2001

Related U.S. Application Data

(63) Continuation-in-part of application No. 09/615,295, filed on Jul. 13, 2000.

(30) Foreign Application Priority Data

May 12, 2000 (GB) 0011543.6

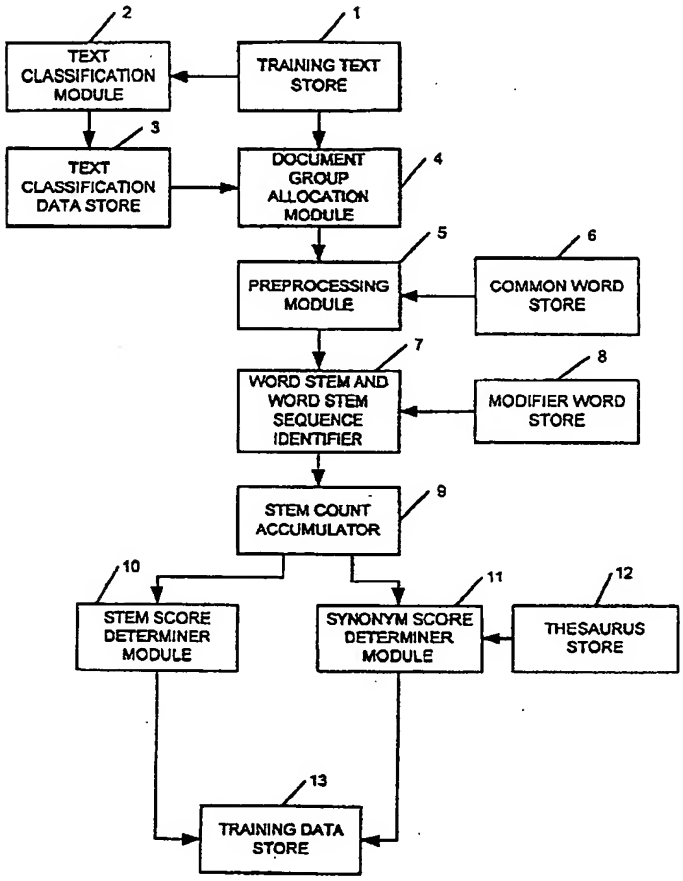
Publication Classification

(51) Int. Cl.[7] G06F 7/00

(52) U.S. Cl. 707/500

(57) ABSTRACT

An automatic text classification system is provided which extracts words and word sequences from a text or texts to be analyzed. The extracted words and word sequences are compared with training data comprising words and word sequences together with a measure of probability with respect to the plurality of qualities. Each plurality of qualities may be represented by an axis whose two end points correspond to mutually exclusive characteristics. Based on the comparison, the texts to be analyzed are then classified in terms of the plurality of qualities. In addition, a fuzzy logic retrieval system and a system for generating the training data are provided.



PGPUB-DOCUMENT-NUMBER: 20020099730

PGPUB-FILING-TYPE: new

DOCUMENT-IDENTIFIER: US 20020099730 A1

TITLE: Automatic text classification system

PUBLICATION-DATE: July 25, 2002

INVENTOR-INFORMATION:

NAME	CITY	STATE	COUNTRY	RULE-47
Brown, Daniel	London		GB	
Janes, Benjamin Anthony	Surrey		GB	
Steele, Murray	London		GB	
Cooper, Richard James	London		GB	

US-CL-CURRENT: 715/500

ABSTRACT:

An automatic text classification system is provided which extracts words and word sequences from a text or texts to be analyzed. The extracted words and word sequences are compared with training data comprising words and word sequences together with a measure of probability with respect to the plurality of qualities. Each plurality of qualities may be represented by an axis whose two end points correspond to mutually exclusive characteristics. Based on the comparison, the texts to be analyzed are then classified in terms of the plurality of qualities. In addition, a fuzzy logic retrieval system and a system for generating the training data are provided.

---------- KWIC ---------

Summary of Invention Paragraph - BSTX (9):

[0009] U.S. Pat. No. 5,781,879 discloses a system for the semantic analysis and modification of information in the form of text. A predetermined lexicon has scores for lexical units (words or phrases) for various categories. Each lexical unit has meaning and semantic content of it's own. The lexicon is used to lookup and accumulate an aggregate score for text for each category. A user is able to modify the text to modify the semantic content of the text by referring the aggregate scores and trying to modify them to preferred values by replacing lexical units in the text with lexical units having different scores for the categories. This system requires a predetermined lexicon having predetermined scores for lexical units for the categories. Each category is given a discrete score and a score is assigned for each category only for individual lexical units. Thus the accumulated score is accumulated using only discrete values for single lexical units and does not provide a system that uses rich semantic information in the text and in training texts.

Detail Description Paragraph - DETX (125):

[0162] FIG. 13 illustrates the hierarchical structure of a classification tree in accordance with an embodiment of the present invention. In this embodiment the qualities or axes have extreme values indicating how much the document is concerned with a topic such as Money. Thus the extremes can be simply YES and NO. This hierarchical structure requires 4 classifiers having 4 different sets of training data. In this embodiment the documents are all from the Reuters news feed. A first set of training data and a first classifier will thus provide 3 qualities or axes for which the documents are given scores by automatic or manual classification. The word stems and word stem sequences in the documents are identified to obtain the training data which will give scores for the 3 axes: Grain, Money and Crude and the associated distribution of word stem and word stem sequence scores across the groups as described above and as illustrated in FIG. 7. A second set of training data and a second classifier will provide 2 qualities or axes: Corn and Wheat for which a subset of the documents having the highest scores for the Grain classification are given scores by automatic or manual classification. The word stems and word stem sequences in the subset of documents are identified to obtain the training data which will give scores for the 2 axes: Corn and Wheat and the associated distribution of word stem and word stem sequence scores across the groups as described above and as illustrated in FIG. 7. A third set of training data and a third classifier will provide 2 qualities or axes: Dollar and Interest for which a subset of the documents having the highest scores for the Money classification are given scores by automatic or manual classification. The

word stems and word stem sequences in the subset of documents are identified to obtain the training data which will give scores for the 2 axes: Dollar and Interest and the associate distribution of word stem and word stem sequence scores across the groups as described above and as illustrated in FIG. 7. A fourth set of training data and a fourth classifier will provide 2 qualities or axes: Gas and Shipping for which a subset of the documents having the highest scores for the Crude classification are given scores by automatic or manual classification. The word stems and word stem sequences in the subset of documents are identified to obtain the training data which will give scores for the 2 axes: Gas and Shipping and the associated distribution of word stem and word stem sequence scores across the groups as described above and as illustrated in FIG. 7. Thus the highest score for one of the qualities or axes will determine the classification assigned e.g. Money and hence the next set of classifications e.g. Dollar and Interest.